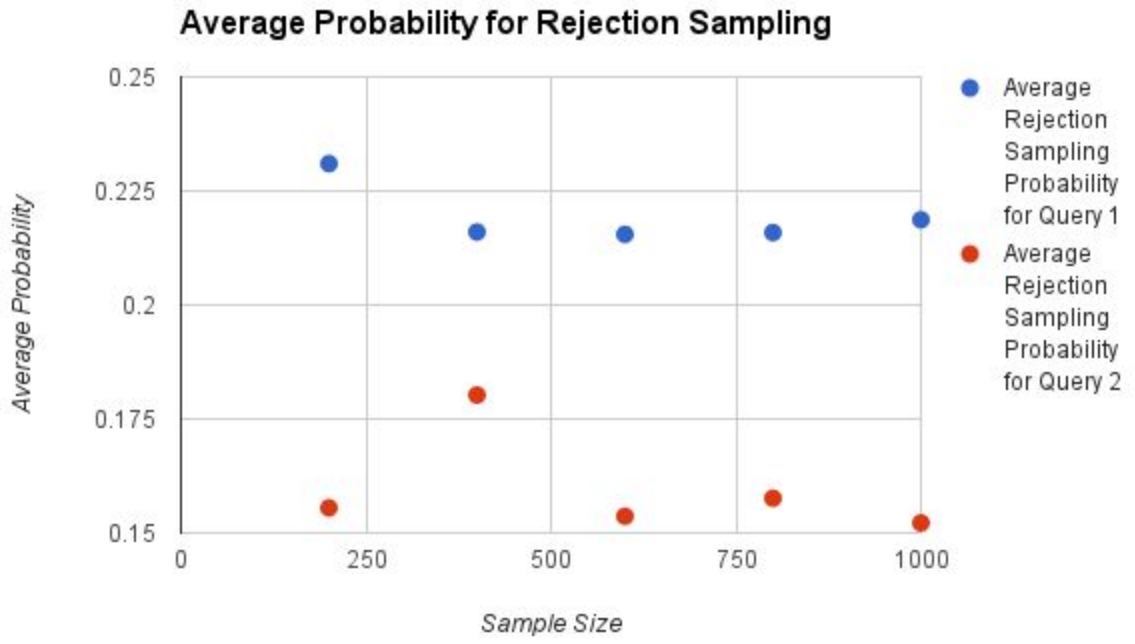
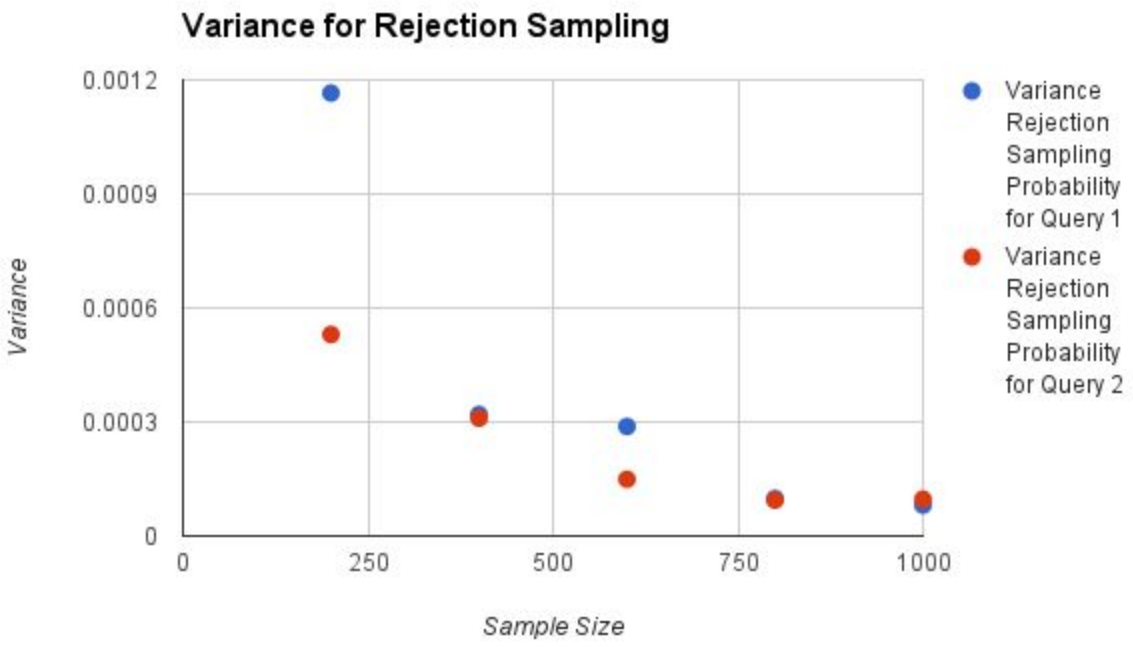
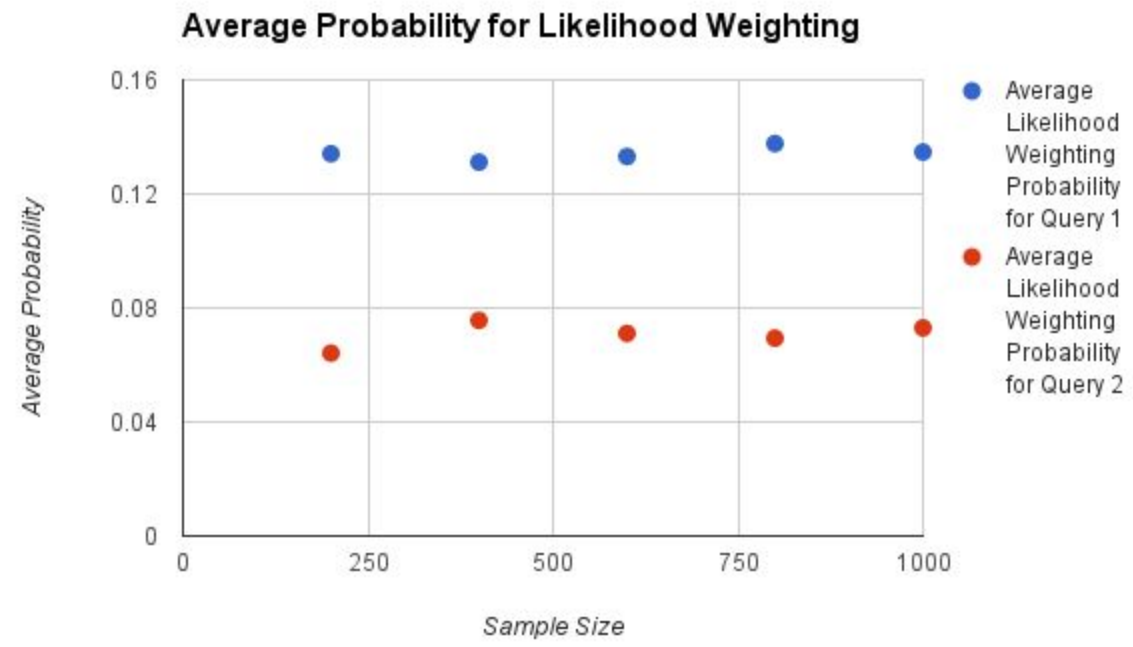
Jacob Hackett
CS4341 A16 Project 6 Writeup

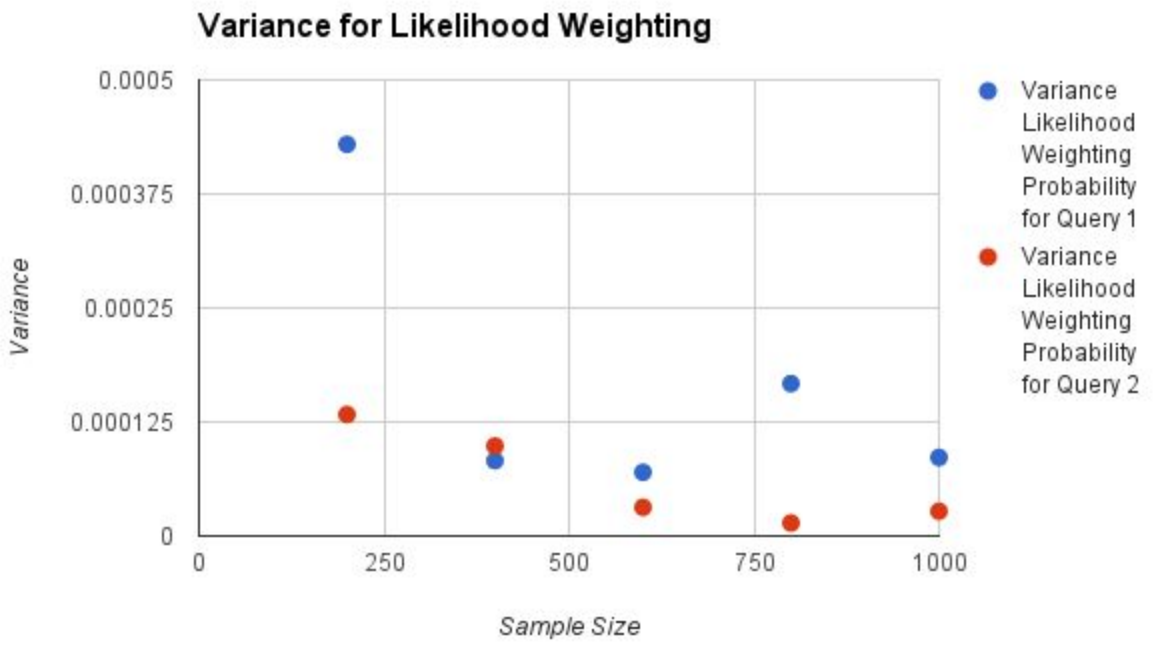
1. **Plot the means of each sampling method against each other for each query file. Plot the variances of each sampling method against each other for each query file.**

This writeup describes the trials I performed on my program. I performed the trial on each of the two provided query files on sample sizes of 200, 400, 600, 800, and 1000. I did this 10 times for each sample size. The congregation of data can be seen on a Google Docs sheet here: https://docs.google.com/spreadsheets/d/1d21uUeOc8BFXa6hu78ENlRyDRU3kvfJ3jT8RL2JpbjY/edit?usp=sharing The results of this data are in the four charts below. These charts are also in the sheet links above.




Variance for Rejection Sampling
0.0012
0.0009
0.0006
0.0003
0
Variance
0
250
500
750
1000
Sample Size
Variance Rejection Sampling Probability for Query 1
Variance Rejection Sampling Probability for Query 2


Average Probability for Likelihood Weighting
0.16
0.12
0.08
0.04
0
Average Probability
0
250
500
750
1000
Sample Size
Average Likelihood Weighting Probability for Query 1
Average Likelihood Weighting Probability for Query 2



2. **Did either method converge to a probability? Was there any difference in the convergence rate? If so, for each case, state which algorithm converged faster and *explain why*.**

For query 1, rejection sampling converged towards roughly 0.219. You can see the rate of convergence was increasing as sample size increased. For query 2, rejection samplying converged towards roughly 0.152. You can see the same behavior regarding rate of convergence in the variance graph.

For query 1, likelihood weighting converged towards roughly 0.135. You can see the rate of converge was increasing by the variance approaches zero as the sample size increases. For query 2, likelihood weighting converged towards 0.073. The variance is also also decreasing and approaching zero suggesting this is approaching the true value as predicted by the algorithm.

The rate of convergence between the two algorithms was different. Likelihood weighting converged faster and the variance was very close to zero by the time I hit 600 samples whereas rejection sampling took much longer and showed the same progress around 1000 samples. This is because likelihood weighting treats the probabilities that are more likely to be true as more important. Therefore, it was able to ignore more of the bad predictions.

3. **Write a readme file that documents any external libraries or design choices you made in your project, as well as the language you used and compilation**

instructions if necessary. Include any information that may be helpful for the TA while grading.

My README.md is in my project submission. Please see it there. Thank you!